DOCUMENT RESUME

ED 338 014 FL 019 185

AUTHOR Kasper, Gabriele; Dahl, Merete
TITLE Research Methods in Interlanguage Pragmatics. Technical Report #1.
INSTITUTION Hawaii Univ., Manoa. Second Language Teaching and Curriculum Center.
SPONS AGENCY Department of Education, Washington, DC.
PUB DATE Mar 91
NOTE 60p.
PUB TYPE Reports - Descriptive (141)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Comparative Analysis; Data Collection; *Interlanguage; *Language Research; *Pragmatics; *Research Methodology; *Second Language Learning

ABSTRACT

Methods of data collection are reviewed that were used in 39 studies of interlanguage pragmatics, defined narrowly as the investigation of non-native speakers' comprehension and production of speech acts, and the acquisition of second language-related speech act knowledge. Data collection instruments are distinguished according to the degree to which they constrain informants' responses, and whether they tap speech act perception/comprehension or production. A main focus of discussion is the validity of different types of data, particularly their adequacy to approximate authentic performance of linguistic action. It is concluded that there is a clear need for more authentic data, collected in the full context of the speech event, and for comparative studies of the validity of different elicitation techniques. Contains approximately 55 references. (Author/LB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED338014

KASPER • DAHL

RESEARCH METHODS IN INTERLANGUAGE PRAGMATICS

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SECOND LANGUAGE TEACHING & CURRICULUM CENTER
University of Hawai'i at Mānoa

FL019185

BEST COPY AVAILABLE

GABRIELE KASPER
University of Hawai'i

MERETE DAHL
University of Copenhagen

RESEARCH METHODS IN INTERLANGUAGE PRAGMATICS

TECHNICAL REPORT #1

March 1991

SECOND LANGUAGE TEACHING & CURRICULUM CENTER
University of Hawai'i at Mānoa

TECHNICAL REPORT #1

RESEARCH METHODS IN INTERLANGUAGE PRAGMATICS

This technical report reviews the methods of data collection employed in 39 studies of interlanguage pragmatics, defined narrowly as the investigation of nonnative speakers' comprehension and production of speech acts, and the acquisition of L2-related speech act knowledge. Data collection instruments are distinguished according to the degree to which they constrain informants' responses, and whether they tap speech act perception/comprehension or production. A main focus of discussion is the validity of different types of data, in particular their adequacy to approximate authentic performance of linguistic action.

Second Language Teaching & Curriculum Center
Webster 203
University of Hawai'i
Honolulu, Hawai'i 96822

Funds for the publication of this technical report were provided in part by a grant to the University of Hawai'i under the National Language Resource Centers program of the U. S. Department of Education.

INTRODUCTION

IT IS NOW well accepted that different tasks constrain language use in different ways. In the study of syntax and semantics, it has predominantly been psychological constraints on language processing that have been found to account for variation dependent on task and task demands (Crookes, 1989; Hulstijn, 1989). In pragmatics, we are dealing with a double layer of variability: (a) variability which reflects the social properties of the speech event, and the strategic, actional and linguistic choices by which interlocutors attempt to reach their communicative goals; and (b) the variability induced by different instruments of data collection. While our primary goal is to uncover sociopragmatic and pragmalinguistic variability "in the real world," we have to be alerted to task effects induced by our instruments in order to assign correct causal interpretations to observed variation. The purpose of this paper is to provide a descriptive overview of methods of data collection used in interlanguage pragmatics, and to evaluate their validity relative to the research questions under study.

In order to delimit the scope of this review, the following criteria will be adopted:

1) *Interlanguage pragmatics* will be defined in a narrow sense, referring to nonnative speakers' (NNS) comprehension and production of speech acts, and how their L2-related speech act knowledge is acquired. Studies addressing conversational management, discourse organization, or sociolinguistic aspects of language use such as choice of address forms will be outside the scope of this paper.

2) As the term *interlanguage* (IL) suggests, the studies to be included are those which examine child or adult NNS speech act behavior and knowledge, to the exclusion of L1 child and adult pragmatics. Even though the comparison of methods used in developmental and adult L1 pragmatics with those in interlanguage pragmatics would be an intriguing issue, we did not consider this a practicable task, in view of the extensive body of research existing in both of the former fields. However, we shall discuss two studies on L1 pragmatics towards the end of the paper, to compensate for the paucity of comparative methodological research in IL pragmatics.

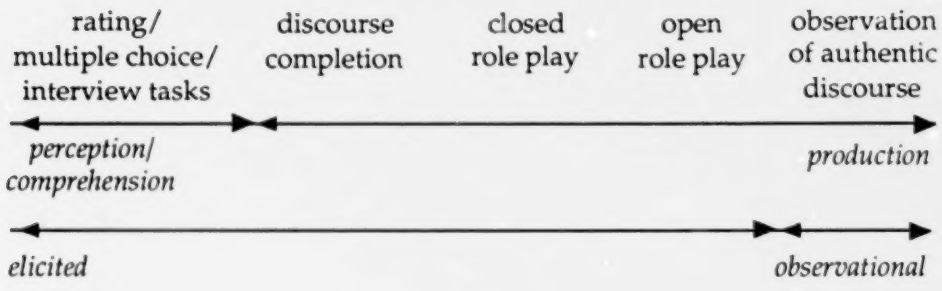
3) The focus will be on methods of data collection rather than analysis. Methods of analysis such as the analytical model, units of analysis, individual categories and statistical treatment have been shown to determine research outcomes (Slethei,1990). Yet in more than one sense, data collection is primary in relation to analysis: not only because it comes prior to analysis in the sequential organization of the research process, but also because it is a more powerful determinant of the final product. Problems with coding and quantification can, in principle, be remedied upon detection; however, if raw data are flawed because the instrument or observation procedure was inadequate, repair is often not feasible, and the value of the study is questionable.

A further caveat is in order. Whereas IL pragmaticists have been concerned about the validity of their data collection procedures, as will become apparent in this review, no tests of the reliability of the deployed data collection instruments have been reported in the literature. Since, according to classical measurement theory, reliability constitutes the upper limit of validity, the lack of information about the reliability of the instruments used constrains claims about their validity (e.g., Kirk & Miller, 1985). It thus remains a requirement for future IL pragmatics research to subject data collection instruments to tests of reliability.

TYPES OF DATA COLLECTION PROCEDURES

THE METHODS of data collection employed in IL pragmatics can be characterized in terms of the constraints they impose on the data: the degree to which the data are predetermined by the instrument, and the modality of language use subjects/informants are engaged in. Figure 1 arranges different methods along these two dimensions (cf. Larsen-Freeman & Long (in press) for a similar display).

Figure 1 : Methods of data elicitation



The procedures towards the left-hand side of the continuum — different kinds of rating tasks (paired comparison, card sorting, rating scales), multiple choice questionnaires, and interviews provide information about subjects' perception of alternative speech act realizations, for example in terms of relative politeness, or about the pragmatic meaning subjects assign to offered stimulus material. The production procedures from the middle to the right-hand side of the continuum comprise highly constrained instruments such as Discourse Completion questionnaires and 'closed' role plays with no interaction, and more weakly controlled 'open' role plays with partially self-directed interaction between the players. The right-most side represents observational studies where no deliberate constraints are imposed on the informants, although there may be unintentional observer effects.

Table 1 : Methods eliciting perception and comprehension data

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln	Instrument
Walters 1979	requests	intermediate/ advanced	75	English	diverse (17)	60 10	English Spanish	paired comparison
Carrell & Konneker 1981	requests	intermediate/ advanced	73	English	diverse	42	English	card sorting
Tanaka & Kawade 1982	requests	advanced	(1) 10 (2) 32	English English	Japanese diverse	10 53	English English	card sorting, multiple choice
Olshtain & Blum-Kulka 1985	requests, apologies	not reported	124	Hebrew	not reported	160 172	Hebrew English	3 point scale
Carrell 1979	indirect answers	intermediate/ advanced	72	English	diverse	29	English	multiple choice
Carrell 1981b	requests	low-int./int. high-int./adv.	82	English	diverse	not reported	English	multiple choice
Kasper 1984	responding acts	intermediate/ advanced	48 pairs NNS-NS	English	German	48 pairs 48 pairs	English German	open role play

Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

METHODS ELICITING PERCEPTION AND PRODUCTION DATA

MOST OF THE EARLY STUDIES of IL pragmatics probed into NNS' perception and comprehension of speech acts carried out in L2 (see Table 1).

Walters (1979) examined the perception of politeness in fourteen generic request strategies through card paired comparison, such as deciding on the relative politeness value of "shut up" versus "please be quiet." Subjects were explicitly instructed to ignore context as much as possible. Even though English NS and NNS of various language backgrounds agreed in their overall perceptions of politeness, the NNS (and female NS) were more categorical in their judgments, and displayed greater unanimity than the male NS.

Carrell and Konneker (1981) looked at NNS' perceptions of politeness in eight request strategies. Subjects were presented with cards specifying different request contexts and the eight strategies, and asked to sort the strategies according to politeness. The rank order of politeness values obtained for each strategy suggested that NNS both overdifferentiate request strategies (they perceived seven politeness levels, whereas the NS distinguished only five) and underdifferentiate strategies (they did not recognize boundaries between strategies where NS did).

These findings were corroborated by Tanaka and Kawade (1982), who replicated Carrell and Konneker's (1981) study with native Japanese subjects. In a second study reported in the same paper, Tanaka and Kawade investigated NNS's perceptions of appropriate politeness strategies dependent on social context. The instrument was a questionnaire including pictures and descriptions of situations in which a person wished to borrow something, and a list of six requestive strategies from which subjects had to choose the one they perceived as most adequate in the context. While both NS and NNS selected more polite strategies with increased social distance, NNS overall chose less polite strategies than the NS.

Olshtain and Blum-Kulka (1985) studied the perception of politeness in requests and apologies by NNS of Hebrew. A questionnaire including four request and four apology situations, followed by six request

and apology strategies, respectively, was administered to subjects, who had to rate each strategy for politeness on a 3-point scale. The effect which accounted most strongly for differences between NS' and NNS' politeness perceptions proved to be NNS' length of stay in the target community. Over time, the NNS' tolerance for more directness and positive politeness (in the sense of Brown & Levinson, 1987) increased, indicating acculturation to the target culture.

The studies of NNS' perception of politeness in speech act strategies can be described as metapragmatic judgment studies: the issue in focus is not on-line perception of politeness in context but relatively permanent states of pragmatic knowledge. It is therefore legitimate to probe into the politeness values assigned to "generic" speech act realization strategies (see Walters, 1979), since these can be assumed to be cognitively represented as prototypical realization patterns. This research goal involves considerable methodological problems, however. As Ellis (in press) points out, competence is not available for immediate inspection. Judgment data are a special kind of performance data, the interpretation of which has to take into account effects that are difficult to control for. Among these effects are subjects' subjective understanding of the task, and effects deriving from the semantic content and context. It seems that pragmaticists in particular are up against a no-win situation: if utterances to be judged for politeness or illocutionary force are stripped of content and context, subjects are likely to supply some of this information anyway and base their judgments on the mentally elaborated versions – instructions to the contrary (Walters, 1979) cannot preclude this. If context is provided, as in Olshtain and Blum-Kulka (1985), further subjective elaboration is still possible, and research findings are harder to generalize to contexts with different properties. What is needed here, clearly, are carefully conducted studies that compare effects of different task conditions in the study of metapragmatic knowledge. We are not aware of any such studies.

Carrell (1979) studied NNS' comprehension of indirect answers, such as in:

A: Did you scrub the floor?
B: Well I swept it.

Her instrument was a written questionnaire with twenty-seven short dialogues followed by three multiple choice answers. For example:

> Bob comes up to Ann in the student center. Bob says: "Did you go to the movies last night?" Ann says: "I had to study last night."
>
> a) Ann went to the movies last night.
> b) Ann didn't go to the movies last night.
> c) No idea whether (a) or (b).

Overall the NNS did well in inferring the intended meaning from the input, but they had problems interpreting indirect answers such as in "A: Is Greg smart? – B: Is Sadat Jewish?"

In her 1981b study, Carrell examined NNS' comprehension of request forms, using a procedure developed by Clark and Lucy (1975) and replicated by Carrell (1981a) with NS adults and children. The data from the 1981a study constituted the baseline for the study at hand. Subjects listened to forty tape-recorded requests that asked the listener to color a circle red, or not to color a circle red. The experimenter then showed them a blue or a red circle, upon which subjects had to decide whether the requester would like the experimenter's choice of colored circle, and mark their decision on an answer sheet. From the correctness of responses, a developmental pattern emerged: positively conveyed requests were more easily understood than their negative counterparts (e.g., the circle needs/doesn't need to be colored red). The most important determinant of ease or difficulty of request comprehension for all groups of subjects, irrespective of age, native-nonnativeness and NNS' proficiency level, was found to be the linguistic properties of the request.

Carrell's two comprehension studies (1981a,b) examined the on-line attribution of illocutionary force to incoming utterances. Unlike Clark and Lucy in their original study (1975), she used response accuracy rather than reaction time as a measure, and she did not make any claims about the

sequential attribution of force (i.e., whether literal meaning is assigned before indirect meaning). Yet the validity problem raised in discussions of the original study obtains in Carrell (1981a,b) as well: as Gibbs (e.g., 1981) has demonstrated in a series of studies, assignment of force in on-line processing depends on the conventionality of both the request and its context. It is a research issue not yet addressed in IL pragmatics as to whether NNS' processing modes exhibit the same sensitivity to context and conventionality as NS', and if their processing styles change developmentally (but cf. Takahashi, forthcoming).

Whereas the studies reported above examined perception and comprehension through subjects' reaction to stimulus material specifically designed for the purpose, Kasper (1984) deployed conversational performance data to infer learners' pragmatic comprehension from their responses to an interlocutor's preceding turn. The data were elicited by means of open-ended role plays performed by 48 dyads of German learners of English interacting with English NS, and by the same number of German and English NS dyads. Learners' second-pair parts often lacked coherence with the preceding first-pair part, either on the referential level, the relational level, or both. The observed lack of coherence was attributed to excessive use of bottom-up processing which made it difficult for learners to adopt adequate inferencing strategies at the discourse level (cf. also Stemmer, 1981).

By studying pragmatic (mis-)comprehension in ongoing discourse, the possible impact of the conversational history (as opposed to noninteractional contextual information as in, e.g., Carrell 1979, 1981a,b) on the listener's response can be assessed. However, a distinct disadvantage of the procedure is that due to the open-endedness of the role play data, instances of contextually inappropriate responses can be very low. Furthermore, the high-inference quality of the data makes the comprehension processes underlying the responses difficult to reconstruct: after all, particularly in the case of NNS, inappropriate responses may well be due to production problems rather than to misunderstanding. This data type seems therefore more suitable for an initial exploratory investigation

than for a hypothesis-testing study. One possibility of increasing the accuracy of the investigator's analysis would be to consult subjects' own perceptions of their interlocutors' contributions, and their responses to them, via retrospective verbal reports (e.g., Faerch & Kasper, 1987). Additionally, subsequent studies employing more controlled data types could be conducted (cf. Grotjahn, this issue, on Subjective Theories).

METHODS ELICITING PRODUCTION DATA

DISCOURSE COMPLETION

Discourse Completion tasks have been a much used and much beleaguered elicitation format in cross-cultural and IL pragmatics (see Table 2). The format was first developed by Levenston and Blum (1978) to study lexical simplification, and first adapted to investigate speech act realization by Blum-Kulka (1982). Discourse Completion tasks are written questionnaires including a number of brief situational descriptions, followed by a short dialogue with an empty slot for the speech act under study. Subjects are asked to fill in a response which they think fits the given context.

In her study of request realizations in Hebrew, Blum-Kulka (1982) found that the NNS used the same range of strategies as the NS, yet their strategy choice differed in contextual distribution. The NNS preferred less direct requestive strategies than the NS.

A series of studies based on the same Discourse Completion questionnaire was conducted in the Cross-Cultural Speech Act Realization Project (CCSARP, cf. Blum-Kulka & Olshtain, 1984; Blum-Kulka, House, & Kasper, 1989). The questionnaire consisted of eight request and eight apology contexts, and was translated into a variety of languages.

Blum-Kulka and Olshtain (1986) investigated length of utterance in the request performance in Hebrew of NS and NNS at three levels of proficiency. They found that utterance length varied with proficiency, the high intermediate learners using more long-winded request realizations than the NS and the low intermediate and advanced learners. The

Table 2 : Production - Discourse completion model

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln	Items
Blum-Kulka 1982	requests	intermediate/ advanced	44	Hebrew	not reported	32 10	Hebrew English	17
Blum-Kulka & Olshtain 1986	requests	low-intermediate/ high-intermediate/ advanced	240	Hebrew	English (142)	172	Hebrew	5
House & Kasper 1987	requests	intermediate/ advanced	200 200	English English	German Danish	200 163 100	German Danish English	5
Faerch & Kasper 1989 (also Kasper 1989)	requests	intermediate/ advanced	200 200	English German	Danish Danish	100 163	English Danish	5
Svanes in press	requests	beginning/ intermediate/ advanced	60	Norwegian	diverse	148	Norwegian	5
Olshtain & Weinbach 1987	complaints	intermediate/ advanced	35	Hebrew	not reported	35	Hebrew	20

Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

Table 2 : Production - Discourse completion model (continued)

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln	Items
Takahashi & Beebe in press (also Beebe, Takahashi & Uliss-Weltz 1990)	refusals	undergrad/ grad	20 EFL 20 ESL	English	Japanese	20 20	English Japanese	12
Takahashi & Beebe in press	corrections	intermediate/ advanced	15	Englsih	Japanese	25 15	English Japanese	2
Banerjee & Carrell 1988	suggestions	advanced	28	English	Chinese Malay	12	English	60

Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

developmental effect was attributed to paucity of linguistic knowledge in the case of the lowest proficiency group, and adjustment to NS norms in the highest group. The high intermediate learners' verbosity was assessed in communicative effect as a violation of the maxim of quantity, and explained in terms of a playing-it-safe strategy: learners with a high, yet still nonnativelike proficiency increase their verbal output to ensure that they are understood.

Using the same Discourse Completion questionnaire, House and Kasper (1987) compared the request realizations of NNS with different L1s (German and Danish) learning the same L2 (English), while Faerch and Kasper (1989) examined the request strategies used by NNS with the same L1 (Danish) in two different L2s (English and German). The studies demonstrated the NNS' contextual sensitivity to choice of directness levels, which in most cases was consistent with L2 subjects' preferences. Differences occurred in the NNS' selection of syntactic and lexical downgraders, which were used with less frequency and variety than NS', and were partially influenced by L1 transfer. The studies further corroborated Blum-Kulka and Olshtain's (1986) "too many words" finding: all NNS groups displayed more supportive moves (e.g. justifications of their requests) than the target NS (cf. also Kasper, 1989).

Svanes (in press) used the CCSARP Discourse Completion questionnaire to study request realizations by NNS of Norwegian with different L1s. Her overall results agreed with those in House and Kasper (1987). No significant differences were established between the NNS groups.

Olshtain and Weinbach (1987) examined complaining strategies used by NNS and NS of Hebrew. Every type of strategy was used by both groups, with a preference for either mild or strong complaints. Overall the NNS opted for milder, the NS for stronger complaint strategies.

In two studies, Beebe and collaborators investigated refusals by Japanese learners of English (Takahashi & Beebe, 1987; Beebe, Takahashi, & Uliss-Weltz, 1990). The Discourse Completion items comprised four stimulus types – requests, invitations, offers, and suggestions – and were

varied according to addressee status: equal, higher and lower. L1 transfer was evident in the type, order, and frequency of the semantic formulae used by the NNS. In Takahashi and Beebe (1987), where the variables studied included learners' proficiency and EFL versus ESL, proficiency effects were not clearly identifiable, although the EFL learners displayed more transfer.

Using a 12-item questionnaire including 2 items on corrections, Takahashi and Beebe (in press) examined how the speech act of correction was performed by Japanese NNS of English as compared to NS of English and Japanese. The issues under study were (a) the use of formulae which reduce the face-threatening impact of corrections, such as positive remarks, softeners, and expressions of regret; (b) L1 transfer in the Japanese learners' responses; and (c) style shifting according to interlocutor status. The decisive variables determining the choice of mitigating formulae and style shift were found to be the status relationship between the interlocutors – the lower status speakers opting out more, using more softening formulae but avoiding positive remarks – and the type of situation. Transfer from Japanese patterns of style-shifting and use of softeners were also identified in the NNS responses.

Banerjee and Carrell (1988) studied the performance of suggestions by Chinese and Malay NNS of English. Questionnaire items were systematically varied for sex, familiarity, and degree of embarrassment to the addressee. Compared to NS controls, the NNS were less likely to make suggestions in embarrassing or slightly embarrassing situations, yet they agreed with NS in choosing less direct strategies in slightly embarrassing contexts. While the strategy preferred by both groups was a statement of fact, differences were observed in the number and variety of redressive acts. Below, studies employing Discourse Completion questionnaires in conjunction with other measures will be discussed. A preliminary comparison of the studies reported so far point to the following methodological differences:

1) Whereas all studies use L2 baseline data, they differ according to

presence or absence of L1 controls. Whether or not L1 controls need to be included would obviously depend on the specific research questions under study, yet absence of L1 controls precludes examining observed variation for transfer effects. As all investigations using L1 controls report at least some transfer, the canonical design for interlanguage studies — comparable sets of IL, L1, and L2 data – is more informative, and thus preferable, for the study of IL pragmatics. While this methodological requirement applies to any procedure, the reason we emphasize it at this point is that it is not only desirable but quite feasible with Discourse Completion questionnaires: researchers have fewer excuses for *not* using it with Discourse Completion than with data collection methods demanding more resources.

This having been said, a word of caution seems in order about the use of the term *transfer* in IL pragmatics studies. No attempts have been made in the studies reviewed in this paper to define the concept; yet there appears to be some implicit consensus to categorize as transfer any use by NNS of speech act realization strategies or linguistic means which is different from L2 NS use and similar to L1 NS use. This view of transfer as a product rather than a process, dating back conceptually to Lado (1957) and in its operationalized form to Selinker (1969), has long been obsolete in the more comprehensive, process-oriented study of cross-linguistic influence in other linguistic areas (e.g., Kellerman & Sharwood Smith, 1986). Clarifying the concept of pragmatic transfer should have high priority on the research agenda in IL pragmatics.

2) Studies differ according to number of items per questionnaire and number of subjects. An inverse relationship is sometimes observable between these two features: In most of the CCSARP studies, only five Discourse Completion items were included in the request analyses, while the number of subjects ranged from 163 to 240. In Banerjee and Carrell (1988), on the other hand, the Discourse Completion questionnaire consisted of sixty items, and was administered to only 28 NNS and 12 NS. There can be no hard-and-fast rule about number of items and subjects; the more variables studied in each, the higher the requisite number. If, as in Takahashi and Beebe (1987), four different stimulus types and three

different status relationships are to be examined, then a 12-item questionnaire is clearly insufficient, as Beebe, Takahashi, and Uliss-Weltz (1990, p. 67) note. The same problem applies to the CCSARP request studies, where social distance and social power were systematically varied. However, other variables that demonstrably have an impact on request realization, such as requestive goal, degree of imposition, and interlocutors' rights and obligations, were neither systematically varied nor controlled for. Banerjee and Carrell's (1988) questionnaire manages to achieve a systematic variation of context-internal and context-external variables. The reservation we have here is a possible effect of the number of items on the quality and completeness of the responses. In our experience, it often takes considerable tenacity to persuade subjects to complete a 20-item questionnaire. Unlike psychology students, subjects participating in IL studies usually receive neither money nor credits, and what they might perceive as an unreasonable imposition on their time is likely to be reflected in their responses (or lack of same). One might imagine that such effects would multiply with as many as sixty items. In fact, for educational research, Wolf suggests that a "full questionnaire should require certainly less than 30 minutes to complete, and preferably, less than 15 or 20" (1988, p. 481).

The necessary number of subjects would depend both on the subject-related variables under study, and on the number of coding categories. The huge numbers of subjects in the CCSARP data appear somewhat excessive in view of the fact that the samples were homogeneous in terms of age, social stratum, years of study, and L1, so that groups were still large enough when samples were subdivided according to sex. The advantage was that a fine-grained coding schema was applicable to the data. Yet the CCSARP studies consistently show that responses tended to cluster in a few subcategories (such as all subjects' distinct preference for grounders [justifications] as supportive moves for their requests). This tendency is corroborated in an ongoing project on apologies in Thai-English and Japanese-English interlanguage, in which 20-item Discourse Completion questionnaires were administered to 30 or more subjects in each

Table 3 : Production - Role play data

Study	Speech Act	Proficiency	NNS		IL	L1	NS	L2 Ln	Type of Role Play
Walters 1980	requests	not reported*	(1)	10	English	Armenian	10	English	closed role play (puppets)
			(2)				32	Spanish	
			(3)	8	English	diverse	16	English	
			(4)				23	Chicano	
							24	English	
Scarcella 1979	invitations, requests	beginning/ advanced		10 10	English	Arabic	6	English	open role play
Kasper 1981 (also Edmondson, House, Kasper & Stemmer 1984; Stemmer 1981)	initiating (5), responding (6)	intermediate/ advanced	48 pairs NNS-NS		English	German	48 pairs 48 pairs	English German	open role play
Trosborg 1987	apologies	intermediate/ low-advanced/ high-advanced		12 12 12	English	Danish	12 12	English Danish	open role play
Tanaka 1988	requests			4	English NNS-NS	Japanese	4 pairs	English	open role play

* *Subjects in these studies were children age 6 - 15:6.*

Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

gender and language group (Bergman & Kasper, in press) and proficiency level (Maeshiba, Yoshinaga, & Kasper, forthcoming). We feel therefore justified in suggesting that for most purposes in research on IL speech act realization, Discourse Completion questionnaires with 20 items and 30 subjects per undivided sample will serve as a rough guide for decisions on these issues.

ROLE PLAY

Walters (1980) is the only study we are aware of which examines children's IL productive speech act behavior (see Table 3). His data collection technique – children interacting with puppets – is frequently used in L1 developmental pragmatics (e.g., Andersen, 1989). The children were presented with puppets and asked to assist one of the puppets in addressing a request to another puppet. The puppets differed on variables such as age, sex, and race. Results supported the hypothesis that grammatical and pragmatic knowledge are relatively independent abilities: the children used appropriate politeness levels but incorrect forms (e.g., can you know where is the can opener?); false starts and self-corrections indicated that pragmatic and semantic goals were encoded prior to grammatical specification (e.g., will you/can we get in front of you?). Of the studied groups, the Anglo junior high students were most polite (they were also the oldest group), with Armenian and Chicano children being the least polite.

Studies employing closed role plays in combination with supplementary techniques will be reviewed below. Videotaped open role plays were used by Scarcella (1979) in her study of invitations and requests by beginning and advanced Arabic learners of English. There were L2 NS controls but no L1 NS controls. The full conversations were videotaped. It is unclear from Scarcella's report who the subjects' interlocutors were (NS? NNS?). A developmental pattern of politeness markers was discernible, early-acquired forms including "sorry" and "please," while late forms comprised slang, ellipsis, and inclusive "we." The acquisition of politeness forms was found to precede the acquisition of their social distribution.

Kasper (1981) investigated the performance of five initiating speech acts (requests, suggestions, offers, invitations, complaints) and six responding speech acts (acceptances, promises, objections, rejections, apologies, thanks) in audiotaped role plays, carried out by 48 dyads of German learners of English interacting with English NS, and by the same number of L1 and L2 controls. The NNS were successful in reaching referential and actional communicative goals yet mostly unsuccessful in reaching interpersonal goals. Their speech act performance displayed a politeness pattern which was related to neither L1 nor L2. Unlike the NS controls, who behaved according to current politeness theory in that they aggravated acts supportive of hearer's face, and mitigated acts threatening to hearer's face, the NNS tended either not to mark the relational dimension at all, or to mitigate face-supportive acts. As the NNS were EFL learners, it was hypothesized that the *modality reduction* exhibited was induced by years of foreign language teaching, in which interpersonal aspects of communication were largely ignored (cf. also Edmondson, House, Kasper, & Stemmer, 1984). Trosborg (1987) studied the performance of apologies by Danish learners of English at three proficiency levels, interacting with English NS. Compared to L1 and L2 controls, all NNS groups used fewer explanations and minimizing strategies, a fact which was attributed to L1 transfer. With increasing proficiency, the NNS used more modality markers, thus increasing the politeness of their apologies.

Tanaka (1988) examined the request performance of Japanese ESL students interacting with NS friends or lecturers. The scenarios were thus varied according to social distance and social power. No NS controls were used. Throughout the videotaped conversations, the NNS displayed less tentativeness and higher levels of directness than the NS, which the author ascribed to linguistic difficulty: complying with the Cooperative Principle requires less linguistic complexity than being polite. The lack of variation in NNS' requestive behavior in the +social distance and +social power dyads was assumed to reflect Japanese learners' inaccurate perceptions of English interaction, which they seemed to view as more egalitarian and direct than it in fact is.

Open role plays have the advantage that they allow examination of speech act behavior in its full discourse context. Kasper (1981) and Tanaka (1988) have shown, for instance, how request performance can be strategically planned right from the beginning of the conversation, manifesting itself in invested face work and *steering moves* that direct the course of the conversation towards the requestor's goal. Because instructions to accompany open role plays specify players' roles, the initial situation and at least one player's communicative goal, but do not prescribe conversational outcomes, nor how such outcomes are reached, the ensuing interaction is "real" in the context of the play since some outcome needs to be negotiated. Deferring the validity issue to later discussion, it is clear that in comparison to Discourse Completion data, open role plays provide a much richer data source. They represent oral production, full operation of the turn-taking mechanism, impromptu planning decisions contingent on interlocutor input, and hence negotiation of global and local goals, including *negotiation of meaning* (in the SLA sense of the term) when required. The intriguing potential of open role plays for the study of IL pragmatics is that they allow us to observe how speech act performance is sequentially organized (e.g. in terms of strategy choice and politeness investment), what kinds of interlocutor responses are elicited by specific strategic choices, and how such responses in turn determine the speaker's next move. All these are, of course, qualities of authentic conversation (by which we do not mean to say that the enactment of these behaviors is the same under role play and authentic conditions). Role plays have the advantage over authentic conversation that they are replicable and, just as Discourse Completion, allow for the comparative study of NNS and L1 and L2 NS controls. Most role play studies reported in this survey include L1 data. Scarcella (1979) and Tanaka (1988) don't, even though their NNS have the same L1 backgrounds (Arabic and Japanese, respectively). Both for the study of developmental IL pragmatics (Scarcella, 1979) and NNS' use of pragmatic knowledge (Tanaka, 1988), reference to L1 data would seem preferable.

Table 4 : Observation of authentic speech

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln
Wolfson 1989	compliments	diverse	not reported	English	diverse	not reported	English
Bartovi-Harlig & Hartford 1990	status congruent/ incongruent acts	advanced	18	English (6)	diverse (3)	7	English

Table 5 : Multiple instruments to study pragmatic comprehension

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln	Instruments
Ervin-Tripp, et al. 1987	requests	not reported*	28	French	English	23 24	French	narrative, observation
Rintell 1984	emotion	beginning/ intermediate/ advanced	77	English	diverse (13)	19	English	multiple choice, 9 point scale

** Subjects in this study were children age 4- 9.*
Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

A disadvantage which open role plays share with authentic conversational data is that they need transcribing. How much time has to be allotted to the transcription of audiotaped conversation would obviously depend on the delicacy of the transcription required by the purpose of the investigation. Yet our own experience and that of a number of expert transcribers we have consulted suggest that transcribing one hour of a reasonably audible tape in ordinary orthography and including temporal variables takes about ten hours (less for a rough transcription, more for a transcription including prosody and measured pauses). Coding open role play data is more difficult than coding data from more tightly controlled tasks, since illocutionary force and the precise function of conversational markers often cannot be unambiguously determined, facts making interrater reliability harder to achieve.

PRODUCTION: OBSERVATION OF AUTHENTIC SPEECH

From the above, it is understandable why observational data from authentic interactions are underrepresented in IL pragmatics. We have only been able to find two studies of IL speech act realization in this category (see Table 4). Wolfson's (1989b) corpus of compliments comprises two sets of data: "an extensive corpusof well over 1000 examples, gathered over the past 2 years," and including NS and NNS compliments and compliment responses; and "ethnographic data collected through observation and recording of naturally occurring speech in everyday interactions in a wide variety of situations" (p. 227). In the report that follows, no information is provided about the discourse contexts and the participants involved. Results do not refer to compliments but only to NNS' compliment responses. NNS were found to have difficulties in choosing appropriate responses. They failed to appreciate the function of compliments as a social lubricant in American culture, especially as a means to initiate conversation. NNS' minimal responses to compliments counteracted this social function.

Bardovi-Harlig and Hartford (1990) investigated status-preserving strategies in academic advising sessions. Thirty-two advising sessions in a

graduate linguistics program were audiotaped in full length. Participants were 3 faculty members and their graduate student advisees (18 NNS, 7 NS). Clear differences emerged in the ways NS and NNS performed status-incongruent acts, such as suggestions and rejections of advice. The NS succeeded in striking a balance between presenting themselves as active and independent individuals yet preserving the status differential between themselves and their advisers by providing suggestions and rejecting advice with substantial portions of redress: they used a wide range of mitigators and formulated their contributions more tentatively. The NNS offered fewer suggestions and were more definite when they did. The NS produced twice as many rejections of adviser's suggestions and were more successful with their rejections than the NNS.

These two observational studies differ distinctly in approach. Wolfson's impressive corpus of compliments has given important insights into the function of complimenting in American middle-class (East Coast?) society. Her bulge hypothesis (Wolfson 1989a, pp. 129ff) constitutes a major challenge to Brown and Levinson's (1987) politeness theory by demonstrating that politeness investment does not increase in a linear fashion with greater social distance and power, but that most politeness is expended in interaction with friends and colleagues rather than with intimates and strangers. Thus most complimenting is found between acquaintances (in this particular segment of American society, we hasten to add as natives of deference politeness cultures). For IL pragmatics, however, the data presented do not provide information about the NNS' language and cultural backgrounds, proficiency level and relationships to their interlocutors. We do not learn what NNS' use of compliments was like, compared to that of NS. Only inappropriate features in their responses being reported. Bardovi-Harlig and Hartford's (1990) research, on the other hand, examined the speech acts under study in the context of the speech event at large. As NS and NNS interacted with the same NS interlocutors and in the same status relationship, comparability of the data was ensured. Neither of the two studies provided L1 NS controls, so that possible transfer effects could not be analyzed. However, apart from practical difficulties in

obtaining such data, it is particularly clear in the case of the advising sessions examined by Bardovi-Harlig & Hartford that we might here be confronting the limits of cross-cultural comparability, since graduate advising sessions, where adviser and advisee negotiate the student's coursework for the new semester, are institution-bound, and hence culture-specific speech events which have no (direct) equivalent in other cultural contexts. For instance, our Japanese informants report that at Japanese universities, advising takes place informally among the students, older students advising younger students. And even in our native cultures (Germany and Denmark), which have more affinity to the American context, advising sessions of the kind reported do not exist.

A way to compensate for the inherent lack of comparability with NNS' L1 pragmatic behavior in culture-specific speech events, or indeed in all situations where comparative data are hard to come by, would be to conduct retrospective interviews with the participants that can shed light on their perceptions of the preceding interaction. This technique has frequently been used in the analysis of gatekeeping encounters (e.g. Fiksdal, 1988; Grotjahn, this issue). Combining different techniques of data collection has in fact been a procedure employed in a variety of IL pragmatic studies, and it is to these that we now turn.

COMBINED METHODS OF DATA COLLECTION

Different methods in the same study can be used for different purposes. One method can be employed to collect the primary source of data, with data collected by means of another method having the subsidiary function of developing the instrument for the primary data collection or helping with the interpretation of the primary data. Or two or more data types may have equivalent status in the study, yielding complementary information on the research questions at hand. This procedure can also be used for explicit comparison of different data collection techniques.

Techniques to complement primary (production) data typically elicit metapragmatic assessments. Two kinds of assessment data can be

distinguished: (a) assessments of contextual factors which are assumed to affect peoples' perception of a speech event, and hence may explain observed speech act realization patterns; and (b) assessments of the linguistic realization modes themselves, for example, in terms of their directness and politeness and their appropriateness in a given context (see the section on perception and comprehension above). Where assessments of contextual factors are elicited, they often serve as independent variables, explaining or predicting the values obtained for the observed speech act realization patterns as dependent variables.

MULTIPLE APPROACHES TO PRAGMATIC COMPREHENSION

We are aware of only two studies deploying two data elicitation techniques to probe pragmatic comprehension (see Table 5). Rintell (1984) examined how NNS perceive expressions of emotion by combining a measure of on-line perception with a metapragmatic judgment task. After listening to taped dialogues, subjects were asked to identify the expressed emotion on an answer sheet and rate its intensity on a scale. No effects were found for age, sex, and intensity scores. The two variables that did determine NNS' perception of emotive force were L1 and proficiency. Chinese subjects' responses differed consistently from those of Arabic and Spanish students, beginners' perceptions contrasted sharply with those of the intermediate to advanced group.

Ervin-Tripp, Strage, Lampert, and Bell (1987) investigated the kinds of information NNS children draw on in their interpretation of requests: What is the relative importance of situation or activity type as opposed to linguistic elements present in the speaker's utterance? The issue at hand was thus very similar to the one addressed in Carrell's (1979, 1981a, b) studies, reported above. In a pilot study with NS of American English, aged 3 to 7, it was established that these children relied heavily on contextual information in their request comprehension. In the case of NNS children, the success of their request comprehension would depend on their inferencing strategies. If they analyzed the linguistic features of the incoming utterance, their restricted linguistic knowledge should put them

at a disadvantage as compared to NS children; however, if they applied pragmatic reasoning, then age and social experience rather than linguistic proficiency should be decisive. Three age groups (4–9) of second-language learners of French and French and English NS controls were tested by means of two techniques: in the experimental condition, children were presented with narratives depicting typical family scenes and involving implicit requests, followed by the interviewer's instruction to tell the end of the story or to explain what a character in the story meant by her/his utterance. In the natural condition, the experimenter hinted at objects of varying situational relevance present in the setting. Results from both conditions supported a *situation-centered* rather than a *language-centered* model of children's pragmatic comprehension. Cooperation increased with age, and even though degree of explicitness improved the likelihood for cooperation, contextual information was enough to ensure compliance in many instances. Furthermore, it was the children's perception of the practical demands of the situation rather than a reconstruction of speaker's intent that determined their response to a request.

PRODUCTION AND METAPRAGMATIC ASSESSMENT

Eisenstein and Bodman (1986) used a Discourse Completion questionnaire to study expressions of gratitude by advanced ESL learners with different language backgrounds, compared to English NS (see Table 6). In informal interviews following the administration of the questionnaire, some NS and NNS subjects expressed unfamiliarity with some of the situations included in the Discourse Completion task. While the NS produced appropriate responses despite unfamiliarity, the NNS gave more IL-specific responses to items with unfamiliar cultural content. At the same time, subjects did worst on items they assessed as common between cultures, such as thanking a host after dinner. Some items were described by the subjects as involving uncomfortable or embarrassing situations, however, this fact was not reflected in the quality of their responses. The authors concluded that ease or difficulty of particular items appears to be due to a combination of factors, such as required linguistic

Table 6 : Combined production and metapragmatic assessment data

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln	Instrument
Einstein & Bodman 1986	thanking	advanced	67	English	diverse	56 (15)	English	discourse completion task (14), informal interview
House 1988	apologies	intermediate/ advanced	200	English	German	200 163 40 37	English	discourse completion task (7), 3 point scale
Bergman & Kasper in press	apologies	intermediate	288	English	Thai	136 30 30 30	English	discourse completion task (20), 5 point scale
Olshtain 1983 (also Olshtain & Cohen 1983)	apologies	intermediate	13 14	Hebrew Hebrew	English Russian	12 12 12	English Russian Hebrew	closed role play
Garcia	apologies	not reported	10	English	Spanish	10	English	open role play

Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

Table 6 : Combined production and metapragmatic assessment data (continued)

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln	Instrument
Takahashi & Dufon 1989	requests	beginning/ intermediate/ advanced	3 pairs 3 pairs 3 pairs 3 pairs NNS-NS	English	Japanese		English Japanese	open role play
Fraser, Rintell & Walters 1980	requests, apologies	not reported	8	English	Spanish	8 8	English Spanish	closed role play 5 point scale
Rintell 1981	requests, suggestions	not reported	16 10	English	Spanish			closed roleplay

Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

complexity and subjects' familiarity with the situation and cultural disorientation.

In her study of apology performance, House (1988) combined a Discourse Completion questionnaire, administered to German NNS of English and NS of English and German, with 3-point rating scales that assessed the weight of context-external and context-internal factors in the situations included in the questionnaire. The contextual variables were dominance and distance, likelihood for the offender to apologize, addressee's expectation of an apology, offender's obligation to apologize, and imposition involved in the apology. The rating task was administered to NS only (different subjects from those involved in the Discourse Completion task). Overall, ratings were very similar between the English and German NS, the only exception being imposition, which received consistently higher ratings from the German than from the English NS. The use of apology formulae was found to correlate positively with obligation, use of intensifiers correlated negatively with dominance, and no relationship could be established between the expressions of responsibility and any of the context factors. The contextual ratings were helpful in explaining the situation-specific use of apology strategies. It may seem a questionable procedure when ratings are performed by NS yet are used to explain NNS responses as well. However, given the cultural similarity between L1 and L2, as expressed in the ratings by the two NS groups, it appears unlikely that the NNS would have rated differently.

In their study of apology performance by Thai NNS of English in comparison with American English and Thai NS controls, Bergman and Kasper (in press) used a 20-item Discourse Completion questionnaire, designed to include independent variables such as social distance and dominance, interlocutors' age and sex, and severity of offense. A metapragmatic assessment questionnaire elicited subjects' perception of the same offense contexts on a variety of context-external and context-internal factors. Analysis of possible relationships between the contextual factors revealed that the context-external factors distance and dominance were not related to context-internal factors such as severity of offense, offender's

obligation to apologize, likelihood of the apology to be accepted, and offender's loss of face. By contrast, of the context-internal factors, severity correlated highly with obligation, likelihood of acceptance, and face-loss, lending support to the hypothesis that "severity of offense is the representative contextual factor in the socio-pragmatic set of apology" (Olshtain, 1989, p.160). Whereas dominance was not found to have any predictive effect on the choice of apology strategies by the three groups, social distance correlated negatively with explicit admission of responsibility: the closer the interlocutors, the more they would be prepared to assume accountability for their offense overtly. The only apology strategy that correlated with context-internal factors in all three groups was apology intensification, which increased with higher obligation to apologize and offender's face-loss. While all language groups selected their apology strategies according to the specific contextual constellations of the offense situations, three strategies – downgrading responsibility, repair offers, and other kinds of redress such as expressing concern for the offended party – were deployed with consistently higher frequency by the NNS, corroborating previous findings that NNS at an intermediate proficiency level tend to provide more "verbal goods" compared to NS (e.g., Blum-Kulka & Olshtain, 1986; Faerch & Kasper, 1989; House, 1988).

Apologies were also studied by Olshtain (1983) and Olshtain and Cohen (1983). NS of Hebrew, English, and Russian, and English and Russian NNS of Hebrew, first performed closed role plays. Eight apology contexts were typed on cards, and subjects were asked to perform the apology orally, addressed to the experimenter in the role of an addressee. Subsequently metapragmatic assessments were elicited by questionnaire on two issues: Did subjects think that speakers of Hebrew apologize more or less than the subjects in their L1, did the subjects feel that NS of Hebrew might apologize differently from a speaker with the subjects' L1 in any of the eight contexts? The English NNS of Hebrew perceived less need to apologize in Hebrew than in English, which was reflected in a decrease in apology strategies in their Hebrew performance as opposed to their native English. The Russian NS had more universal perceptions of apologies, and

actually increased their apology strategies in Hebrew as compared to their L1.

Garcia (1989) studied apologies performed in open-ended role plays by Venezuelan Spanish NNS of English compared with American English NS. The subjects were female, their interlocutor in the role plays male. No Spanish controls were used. After the role plays, subjects were interviewed individually by the researcher in play-back sessions. Oddly, the interviews do not figure at all in the data analysis and discussion. The findings from the role plays are nevertheless important: in apologizing to a host for not having attended his party, the Venezuelan subjects opted for a positive politeness approach, whereas the American subjects used negative politeness strategies. The incongruence between the politeness style expected by the American host and the NNS resulted in miscommunication and disharmony. In accordance with previous suggestions (Scollon & Scollon, 1983), Garcia proposes that deference politeness may be a preferable style to adopt for cross-cultural communication, since it takes account of interlocutors' territorial concerns and reduces imposition.

Takahashi and DuFon (1989) investigated the request strategies used by Japanese beginning, intermediate, and advanced ESL learners in open-ended role plays. The design and baseline data for this study – Japanese and English NS role plays – were adopted from Takahashi (1987). After the role plays, playback interviews in Japanese were conducted with each subject, the function of which was to identify indirect requests and verify the speaker's intention behind ambiguous requestive utterances (hints). Background data about the subjects were elicited by a questionnaire. A developmental pattern was discernable, according to which the Japanese learners proceeded from more indirect request strategies, which were attributed to L1 transfer, to more direct, targetlike requests. The advanced learners formulated their requests more efficiently and were more successful in achieving compliance. The interview proved to be an important supplementary data source in that it brought out different perceptions of request strategies by the NNS and NS subjects.

In their studies of requests and apologies (Fraser, Rintell, &

Walters, 1980) and requests and suggestions (Rintell, 1981), the researchers combined closed role plays with metapragmatic assessments elicited through rating scales. In Fraser et al. (1980), Spanish NNS of English addressed requests and apologies to a NS interlocutor. Comparison with NS English and Spanish controls demonstrated that for all groups, deference increased with addressee's age. Status increased deference in the NNS and the Spanish NS responses, but did so less in the English NS responses. The transcribed requests were subsequently rated for deference on a 5-point scale, the English requests being rated by both NS and NNS. In both sets of English ratings, older addressees received more deferential requests. In the Spanish ratings, more deference was expended on requests to the opposite sex. The NNS did not transfer their Spanish deference perceptions to English: they rated their Spanish requests as more deferential than their English requests. Rintell (1981) used the same combination of methods as Fraser et al. (1980), but her study did not include NS English controls. She found that receiver's age and sex affected the subjects' perception of deference in requests but not in suggestions. The combination of production and metapragmatic assessment data provides an empirical basis for explaining observed patterns of speech act realization and politeness in terms of perceived contextual constraints, and of the pragmatic force and politeness value language users attribute to different linguistic means and strategies. Metapragmatic assessments of contextual factors can provide an important corrective, or confirmation, of the values and weights of contextual factors built into the instrument by the researcher. Such controls are particularly important in cross-cultural studies where the researcher is not a member of one or more of the implied cultures. Metapragmatic assessments of the force and politeness of realization patterns help explain preferences for certain forms and strategies over others, including specific cultural values associated with different styles. In IL pragmatics, a pertinent issue is to determine whether NNS perceive the illocutionary and politeness value of speech act realization patterns differently from NS, as this would be one reason why NNS might mean something different by saying the same things that NS say, or by saying something different but meaning the same thing as NS do.

Table 7 : Different types of production data

Study	Speech Act	Proficiency	NNS	IL	L1	NS	L2 Ln	Instrument
Rintell & Mitchell 1989	requests, apologies	low/ advanced	29 21	English English	not reported	23 14	English English	discourse completion task, closed
Bodman & Einstein 1988	thanking	not reported	40 pairs 24pairs (NNS-NS)	English	not reported	34 pairs	English	open role play discourse completion task, observation
Beebe & Takahashi 1989a	disagreement embarassing information,	advanced	15	English	Japanese	15	English	discourse completion task (12), observation
Beebe & Takahashi 1989b	disagreement, chastisement	high-int./ advanced	15	English	Japanese	15	English	discourse completion task (12), observation
Beebe & Cummings 1985	refusals				English	22		discourse completion task (1), observation
Dahl forthcoming	refusals, disagreement				English	150		open role play, observation

Note: NNS = number of non-native speakers; NS = number of L2 native speakers.

Finally, an important, and yet virtually unaddressed issue would be whether such perceptions change as NNS' performance in the course of IL development changes.

COMPARING DIFFERENT TYPES OF PRODUCTION DATA

SOME OF THE STUDIES to be discussed in this section were conducted with the explicit purpose of comparing data collected through different procedures; for others, the methodological question was more of a side-issue, although the authors do comment on instrument effects (see Table 7). Two investigations of NS' speech act performance will be included, as these were specifically conducted as inquiries into method.

Rintell & Mitchell (1989) compared ESL learners' and English NS' performance of requests and apologies in Discourse Completion items and closed role plays. The main difference between the two procedures was that the Discourse Completion task elicited written responses, the closed role play oral ones. The NNS' productions differed in length in the two versions, the oral responses being much longer due to the use of more supportive moves, hesitation, and recycling. No length effect was found for the NS. Modality had no impact on the range of request and apology strategies used by NS and NNS. However in some request contexts, the frequency of direct strategies was higher in the written than in the oral condition. The authors suggest that it might be more socially legitimate to use higher directness in writing, and that subjects might have felt uncomfortable about using imperatives in face-to-face interaction with the experimenter. We think that this social interaction effect may be induced by the status difference between students and researcher, and might diminish if the subjects' interlocutor was a familiar status equal. Unless this variable is controlled for, we cannot be sure that the observed effect is due to modality of language use. With the apologies, virtually no difference was found between modalities. The authors conclude that Discourse Completion and (closed) role play yield very similar data, which might

reflect "dramatic equidistance" from natural performance data. We agree and would add that the decisive difference to authentic conversational data on the one hand and Discourse Completion and closed role plays on the other hand is that the two elicitation procedures are non-interactive. This might be a more consequential difference than modality alone.

Following up on their previous study (Eisenstein & Bodman, 1986), Bodman and Eisenstein (1988) examined the expression of gratitude by English NS and NNS. Three data types were used: Discourse Completion questionnaires; open-ended role plays performed by NS-NS, NNS-NNS, and NS-NNS dyads; and field notes of naturalistically occurring expressions of gratitude. All three data types yielded the same words and expressions. Yet the data differed in length and complexity. The Discourse Completion data were the shortest and least complex, the authentic data longest and most complex, with the role play data coming in between. The oral data had more restatements of thanks and discussions about the received gift or service. An important outcome of analyzing the role play and authentic data was that thanking is collaboratively enacted, involving the giver as much as the receiver. The NS role plays were often longer than NNS because the NS did not ask for services or favors directly; rather, they would indirectly suggest their needs. Ritual refusals and expressions of reluctance preceded acceptance of offer and thanking. NS, but not NNS, increased the length of the speech event with greater indebtedness. The less fluent NNS were not able to compensate for shorter expressions by conveying appreciativeness prosodically. The NNS performed better in interaction with NS than with other NNS, who were unable to provide the same kind of support for the less competent partner. The authors conclude that all three data sources contribute importantly to the study of speech act behavior, although the richest data are naturalistically occurring interactions.

In order to examine how advanced Japanese ESL learners compare to English NS in the expression of disagreement, embarrassing information, and chastisement, Beebe and Takahashi (1989a,b) used notebook entries of occurrences of the speech acts under study in authentic conversations, and

a Discourse Completion questionnaire. Substantial differences were found in the ways Americans and Japanese carry out face-threatening acts in English. Results did not support the stereotype of pervasive Japanese indirectness and avoidance of disagreement. The Japanese subjects' choice to express disagreement and chastisement directly or indirectly was determined by their status relationship with their interlocutor. In high-to-low situations, the NNS would disagree and chastise bluntly, whereas the NS would opt for mitigation and positive politeness strategies. In expressing embarrassing information, Japanese question strategies where perceived as face-threatening by the English NS even though they were intended as face-supportive. While Beebe and Takahashi conclude that more natural data is needed to substantiate their findings, they also identify problems in this data type:

> [Natural data] are biased by the linguistic preferences of our friends, relatives and associates. They are also biased in favour of short exchanges, because long ones are impossible to get down word for word in a notebook. And they are biased to ones that the researcher finds especially typical, especially atypical, or especially non-native sounding. It is much harder to notice a native-like ESL example than a distinctly non-native one. Moreover, natural data give us lots of examples that are not at all comparable in terms of speakers, hearers and social situations, unless one or two situations are selected, and this poses other limitations. (1989a)

Some of the problems with authentic data, as Beebe and Takahashi point out so succintly, follow from the *type* of observation referred to. The notetaking technique exceeds the capacity not only of the researcher's notebook and motor skills but also of her short-term memory if the purpose is to record how speech acts are performeds cooperatively. Rather than collecting isolated conversational segments, it is preferable to audio- or videorecord complex speech events, and to compare these data with elicited data types.

The final two studies to be reviewed did just this, but did not

include NNS. As the data types discussed are the same as those frequently used in IL pragmatics, we feel encouraged to relax our criterion of "interlanguage studies only" at this point.

Beebe and Cummings (1985), in a well-known yet unpublished paper, compared refusals performed by native English-speaking ESL teachers in response to a single Discourse Completion item with refusals performed during an authentic telephone conversation. All subjects were female, and those called on the telephone were members of New York State TESOL. The stimulus in the Discourse Completion item and the telephone call were the same: Subjects were asked if they would like to volunteer at the TESOL Convention New York which was going to take place a few weeks later. Comparison of the refusals obtained by both methods resulted in a number of important similarities and differences:

1) The telephone conversations were consistently longer than the Discourse Completion responses. One reason is the obvious fact that they require interaction and negotiation which Discourse Completion precludes. More specifically, in the telephone conversations, participants' co-membership in a social network reduced social distance and invited more talk exchanges than would otherwise be expected between strangers. The authors hypothesized that length and tone of the response will be influenced if subjects responding to a Discourse Completion questionnaire imagine that they are interacting with a familiar interlocutor.

2) In both data types, direct refusals were virtually absent, and refusal strategies were used with the same frequency. The content expressed by the Discourse Completion responses thus matched the content of the conversations. Yet the range of strategies used in the telephone conversations was considerably wider than in the Discourse Completion responses.

3) Differences between the Discourse Completion task and the conversations were most marked in the psychological domain. Some of the telephone conversationalists displayed strong

emotions which would not be called upon in a fictitious situation.

Beebe and Cummings (1985) conclude that Discourse Completion questionnaires are a highly effective means of:

1) gathering a large amount of data quickly;

2) creating an initial classification of semantic formulas and strategies that will occur in natural speech;

3) studying the stereotypical, perceived requirements for a socially appropriate (though not always polite) response;

4) gaining insight into social and psychological factors that are likely to affect speech and performance; and

(5) ascertaining the canonical shape of refusals, apologies, partings, etc., in the minds of the speakers of that language. (1985, pp. 13f.)

Discourse Completion responses do not, however, adequately represent:

1) the actual wording used in real interaction;

2) the range of formulas and strategies used (some, like avoidance, tend to be left out);

3) the length of response or the number of turns it takes to fulfill the function;

4) the depth of emotion that in turn qualitatively affects the tone, content, and form of linguistic performance;

5) the number of repetitions and elaborations that occur; or

6) the actual rate of occurrence of a speech act – e.g., whether or not someone would naturalistically refuse at all in a given situation (Beebe & Cummings, 1985, p. 14).

Dahl (forthcoming) refers to Beebe and Cummings finding that "written role plays bias the response toward less negotiation, less hedging, less repetition, less elaboration, less variety and ultimately less talk" (1985, p. 4). Her research question was whether the same bias would be characteristic of open role plays, and she hypothesized it would not, due to the interactive nature of the data.

Dahl's data comprised three different sets:

1) A sample of 115 taperecorded interviews and 137 corresponding role play interviews, carried out between the researcher and female native speakers of English (students contacted in their dorms). In both types of interview, the researcher asked a student "would you like to do a small role play with me." Thus authentic refusals were collected, and if students agreed to do the role play, the researcher asked the same question again and thus elicited the role play refusals.

2) An authentic group discussion and a role play discussion based on selected contributions to the authentic discussion. The authentic discussion was arranged by Amnesty International and addressed the issue of capital punishment. Based on contributions expressing disagreement and disapproval in the authentic discussion, cue cards for the role play were composed, specifying background situation, divergent opinions previously expressed by other speakers, and the role player's present "intention." Subjects were given leaflets with 6 – 11 numbered cue cards. After familiarizing themselves with the instructions for about 10 minutes, they performed each of the elicited speech acts.

3) Two authentic student discussions, one about capital punishment and one about the relationships between "red" and "green" political strategies (arranged by the local Socialist Workers' Society and the Green Action Society). Fifteen monologic role plays were constructed on the basis of five contributions selected from the authentic role plays. Three female students participated in the role plays.

Despite differences in discourse type, the authentic interactions in the first and second data set exhibited similar features compared to the role plays. In the authentic contexts, informants used more words, more speech acts, more pre-exchanges, took more turns, and produced a higher number of indirect speech acts. No difference was found in the frequency of upgraders and downgraders (aggravators and mitigators), although subcategories were distributed somewhat differently. In the authentic group discussions, informants used more interpersonally oriented upgraders, whereas the role plays displayed more affectually neutral upgraders.

However, comparison of the discussions and monologic role plays in the third data set brought out different results: the monologic role play performances had more words and speech acts than their authentic counterparts. No difference was found in the directness of speech act realization, or the frequencies of upgraders and supportive moves. There were more downgraders, and among them relatively more downgraders expressing positive politeness in the authentic discussions than in the monologic role plays.

The most important features that distinguished between authentic and role play productions across discourse types were amount of talk and directness in the performance of face-threatening acts. Amount of talk also distinguished the two types of role plays from each other, the interactive role plays producing less talk, the monologic role plays more talk than their authentic counterparts.

As amount of talk typically distinguishes between different interlocutor relationships (cf. Wolfson's, 1989a, bulge hypothesis, referred to above), and directness interacts with contextual factors in conveying politeness (see Kasper, 1990, for an overview), the discomforting conclusion suggested by Dahl's study is that role plays are not representative of authentic interaction on these measures.

However, Dahl emphasized that the way the role plays were elicited implied a number of constraints that might have seriously reduced the generalizability of her study. Moreover, she warned that the

circumstances of the data collection might have introduced some extraneous factors which could have impaired the validity of the role plays.

The refusal role plays were enacted immediately after subjects had consented to do a role play. This is a cognitively demanding and potentially confusing task because it requires that subjects take on, and act upon, an attitude exactly opposing the one they displayed just before, and then address the required (and conversationally dispreferred) speech act to the same interlocutor whom they just assured of the contrary. This configuration of cognitive and interactive constraints may well obstruct subjects' access to the schemata they would be able to activate if the interactional content of the role play was more neutral vis-a-vis the embedding real-world context. Despite the familiarity with the situation, then, the lack of role distancing from the immediately preceding interaction will have precluded informants' role identification with their role play part. This lack of opportunity for role identification seems closely related to the specific circumstances of the data collection, and hence constitutes an extraneous effect which could be avoided in differently arranged role plays.

The discussion role plays may have suffered from different interfering effects which nevertheless produced the same results. Subjects were given detailed instructions on which kinds of disagreements and disapprovals to perform, and in which sequence. They were not allowed to recycle arguments already provided. These restrictions imposed high memory load on the subjects and at the same time prevented free negotiation between the interlocutors that would have derived from the interaction itself, such as attending to the interpersonal dimension in performing face-threatening acts.

Dahl (forthcoming) proposes that given the interpersonal, interactional, and cognitive demands made on the subjects, it is perhaps natural for them to make their task manageable by reducing amount of talk and relational concerns, that is, orienting their communicative behavior on the principal of minimal effort. In some of the role play studies reported above (Kasper, 1981; Trosborg, 1987; Tanaka, 1988) this behavior was characteristic of NNS but not of NS subjects. Dahl (forthcoming) therefore warns that the conclusions drawn from her relatively constrained type of

role play cannot be extended to more open and self-directed role play types, and she suggests that the relationship between such open-ended role plays and authentic data should be investigated before any final statements about the validity of role play data are offered.

Rather than representing one well-defined discourse type, role plays comprise a whole range of different tasks, each of which constrains speakers' performance differentially. This is to some extent indicated in Figure 1, yet role plays may be grouped at many different points on an open-closed continuum. Some types of simulations – for instance, where participants retain their own identities – might approximate authentic discourse even more closely than open role plays; yet at the same time the conversational behavior produced may be less predictable, and more difficult to compare both to data elicited by means of the same technique, and to authentic data.

CONCLUSION

THE CONCLUSION seems inescapable: IL pragmaticists are caught between a rock and a hard place. With the exception of highly routinized and standardized speech events, sufficient instances of cross-linguistically and cross-culturally comparable data are difficult to collect through observation of authentic conversation. Conversely, tightly controlled data elicitation techniques might well preclude access to precisely the kinds of conversational and interpersonal phenomena that might shed light on the pragmatics of IL use and development. Clearly there is a great need for more authentic data, collected in the full context of the speech event, and for comparative studies of the validity of different elicitation techniques. Most importantly, the specific contributions of different data collection techniques to different research issues needs further empirical scrutiny. As in all data-based research, a good method is one that is able to shed light on the question(s) under study. Ecological (face) validity should not be a sacred cow in interlanguage pragmatics (nor anywhere else)◆

AUTHORS' NOTE

We are grateful to Marc L. Bergman, Peggy DuFon, Juergen Meisel and Satomi Takahashi for helpful comments on an earlier draft. Special thanks to Eric Kellerman for eliminating the cross-linguistic influence (CLI) in the previous version.

REFERENCES

Andersen, Elaine (1989). *Speaking with style: The sociolinguistic skills of children.* London: Routledge.

Banerjee, Janet & Carrell, Patricia L. (1988). Tuck in your shirt, you squid: Suggestions in ESL. *Language Learning 38*, 313–347.

Bardovi-Harlig, Kathleen & Hartford, Beverly S. (1990). Congruence in native and nonnative conversations: Status balance in the academic advising session. *Language Learning 40:4.*

Beebe, Leslie M. & Cummings, Martha C. (1985, April). *Speech act performance: A function of the data collection procedure?* Paper presented at the TESOL Convention, New York.

Beebe, Leslie M. & Takahashi, Tomoko (1989a). Do you have a bag? Social status and patterned variation in second language acquisition. In Susan Gass, Carolyn Madden, Dennis Preston & Larry Selinker (Eds.), *Variation in second language acquisition: Discourse and pragmatics* (pp. 103–125). Clevedon: Multilingual Matters.

Beebe, Leslie M. & Takahashi, Tomoko (1989b).Sociolinguistic variation in face-threatening speech acts. In Miriam Eisenstein (Ed.), *The dynamic interlanguage* (pp. 199–218). New York: Plenum.

Beebe, Leslie M., Takahashi, Tomoko & Uliss-Weltz, Robin (1990). Pragmatic transfer in ESL refusals. In Robin C. Scarcella, Elaine Andersen & Stephen C. Krashen (Eds.), *On the development of communicative Competence in a second language* (pp. 55–73). New York: Newbury House.

Bergman, Marc L. & Kasper, Gabriele (in press). The interlanguage of apologizing: Cross-cultural evidence. In Shoshana Blum-Kulka & Gabriele Kasper (Eds.), *Interlanguage pragmatics*. New York: Oxford University Press.

Blum-Kulka, Shoshana (1982). Learning how to say what you mean in a second language: A study of speech act performance of learners of Hebrew as a second language. *Applied Linguistics 3*, 29–59.

Blum-Kulka, Shoshana & Olshtain, Elite (1984). Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied Linguistics 5*, 196–213.

Blum-Kulka, Shoshana & Olshtain, Elite (1986). Too many words: Length of utterance and pragmatic failure. *Journal of Pragmatics 8*, 47–61.

Bodman, Jean & Eisenstein, Miriam (1988). May God increase your bounty: The expression of gratitude in English by native and non-native speakers. *Cross Currents 15*, 1–21.

Brown, Penelope & Levinson, Stephen S. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Carrell, Patricia L. (1979). Indirect speech acts in ESL: Indirect answers. In Carlos A. Yorio, Kyle Perkins & Jacquelyn Schachter (Eds.), *On TESOL '79* (pp. 297–307). Washington, D.C.: TESOL.

Carrell, Patricia L. (1981a). Children's understanding of indirect requests: comparing child and adult comprehension. *Journal of Child Language 8*, 329–345.

Carrell, Patricia L. (1981b). Relative difficulty of request forms in L1/L2 comprehension. In Mary Hines & William Rutherford (Eds.), *On TESOL '81* (pp. 141–152). Washington, D.C.: TESOL.

Carrell, Patricia L. & Konneker, Beverly H. (1981). Politeness: Comparing native and nonnative judgments. *Language Learning 31*, 17–31.

Clark, Herbert H. & Lucy, P. (1975). Understanding what is meant from what is said: a study of conventionally conveyed requests. *Journal of Verbal Learning and Verbal Behavior 14*, 56–72.

Crookes, Graham (1989). Planning and interlanguage variation. *Studies in Second Language Acquisition 11*, 367–383.

Dahl, Merete (forthcoming). *Authentic and role play interaction: a study of method*. PhD dissertation, University of Copenhagen.

Edmondson, Willis; House, Juliane; Kasper, Gabriele & Stemmer, Brigitte (1984). Learning the pragmatics of discourse: A project report. *Applied Linguistics 5*, 113–127.

Eisenstein, Miriam & Bodman, Jean W. (1986). 'I very appreciate': Expressions of gratitude by native and non-native speakers of American English. *Applied Linguistics 7*, 167–185.

Ellis, Rod (in press). Grammaticality judgements and second language acquisition. *Studies in Second Language Acquisition 13*,.

Ervin-Tripp, Susan M., Strage, Amy, Lampert, Martin & Bell, Nancy (1987). Understanding requests. *Linguistics 25*, 107–143.

Faerch, Claus & Kasper, Gabriele (Eds.) 1987. *Introspection in second language research*. Clevedon: Multilingual Matters.

Faerch, Claus & Kasper, Gabriele (1989). Internal and external modification in interlanguage request realization. In Shoshana Blum-Kulka, Juliane House & Gabriele Kasper (Eds.), *Cross-Cultural Pragmatics* (pp. 221–247). Norwood, N.J.: Ablex.

Fiksdal, Susan (1988). Verbal and non-verbal strategies of rapport in cross-cultural interviews. *Linguistics and Education 1*, 3–17.

Fraser, Bruce, Rintell, Ellen & Walters, Joel (1980). An approach to conducting research on the acquisition of pragmatic competence in a second language. In Diane Larsen-Freeman (Ed.), *Discourse analysis in second language research* (pp. 75–91). Rowley, Mass.: Newbury House.

Garcia, Carmen (1989). Apologizing in English: Politeness strategies used by native and non-native speakers. *Multilingua 8*, 3–20.

Gibbs, Raymond W. (1981). Your wish is my command: Convention and context in interpreting indirect requests. *Journal of Verbal Learning and Verbal Behavior 20*, 431–444.

Grotjahn, Ruediger (in press). The research program subjective theories: A new approach in second language research. *Studies in Second Language Acquisition 13*.

House, Juliane (1988). "Oh excuse me please...": Apologizing in a foreign language. In Bernhard Kettemann, Peter Bierbaumer, Alwin Fill & Annemarie Karpf (Eds.), *Englisch als Zweitsprache* (pp. 303–327). Tuebingen: Narr.

House, Juliane & Kasper, Gabriele (1987). Interlanguage pragmatics: Requesting in a foreign language. In Wolfgang Loerscher & Rainer Schulze (Eds.), *Perspectives on language in performance: Festschrift for Werner Huellen* (pp. 1250–1288). Tuebingen: Narr.

Hulstijn, Jan H. (1989). A cognitive view of interlanguage variablity. In Miriam Eisenstein (Ed.), *The dynamic interlanguage* (pp. 17–31). New York: Plenum.

Kasper, Gabriele (1981). *Pragmatische Aspekte in der Interimsprache*. Tuebingen: Narr.

Kasper, Gabriele (1984). Pragmatic comprehension in learner native speaker discourse. *Language Learning 34*, 1–20.

Kasper, Gabriele (1989). Variation in interlanguage speech act realization. In Susan Gass, Carolyn Madden, Dennis Preston & Larry Selinker (Eds.), *Variation in second language acquisition: Discourse and pragmatics* (pp. 37–58). Clevedon: Multilingual Matters.

Kasper, Gabriele (1990). Linguistic politeness: Current research issues. *Journal of Pragmatics 14*, 193–218.

Kellerman, Eric & Sharwood Smith, Michael (Eds.) 1986. *Cross-linguistic influence in second language acquisition*. Oxford: Pergamon.

Kirk, Jerome & Miller, Marc L. (1985). *Reliability and validity in qualitative research*. London: Sage.

Lado, Robert (1957). *Linguistics across cultures*. Ann Arbor: University of Michigan.

Larsen-Freeman, Diane & Long, Michael H. (1990) *An introduction to second language acquisition research*. London: Longman.

Levenston, Eddie & Blum, Shoshana (1978). Discourse completion as a technique for studying lexical features of interlanguage. *Working Papers in Bilingualism 15*, 13–21.

Maeshiba, Naoko, Naoko Yoshinaga & Gabriele Kasper (forthcoming). On saying you're sorry in English and Japanese. Ms, University of Hawaii at Manoa.

Olshtain, Elite (1983). Sociocultural competence and language transfer: The case of apology. In Susan Gass and Larry Selinker (Eds.) *Language transfer in language learning* (pp. 232–249). Rowley, Mass.: Newbury House.

Olshtain, Elite & Blum-Kulka, Shoshana (1985). Degree of approximation: Nonnative reactions to native speech act behavior. In Susan M. Gass & Carolyn Madden (Eds.) *Input in second language acquisition* (pp.303–325). Rowley, Mass.: Newbury House.

Olshtain, Elite & Cohen, Andrew (1983). Apology: A speech act set. In Nessa Wolfson & Elliott Judd (Eds.) *Sociolinguistics and second language acquisition* (pp. 18–35). Rowley, Mass.: Newbury House.

Olshtain, Elite & Weinbach, Liora (1987). Complaints: A study of speech act behavior among native and non-native speakers of Hebrew. In Jef Verschueren & Marcella Bertucelli-Papi (Eds.) *The pragmatic perspective* (pp. 195–208). Amsterdam: Benjamins.

Rintell, Ellen (1981). Sociolinguistic variation and pragmatic ability: A look at learners. *International Journal of the Sociology of Language 27*, 11–34.

Rintell, Ellen (1984). But how did you feel about that? The learner's perception of emotion in speech. *Applied Linguistics 5*, 255–264.

Rintell, Ellen & Mitchell, Candace J. (1989). Studying requests and apologies: An inquiry into method. In Shoshana Blum-Kulka, Juliane House & Gabriele Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 248–272). Norwood, N. J.: Ablex.

Scarcella, Robin (1979). On speaking politely in a second language. In Carlos A. Yorio, Kyle Perkins & Jacquelyn Schachter (Eds.) *On TESOL '79* (pp. 275–287). Washington, D.C.: TESOL.

Scollon, R. & S.B.K. Scollon. 1983. Face in interethnic communication. In Jack C. Richards & Richard W. Schmidt (Eds.) *Language and communication* (pp. 156–188). London: Longman.

Selinker, Larry (1969). Language transfer. *General Linguistics 9*, 67–92.

Siethei, Kolbjoern (1990). 'Speech acts', spraakhandlinger og spraakforskning. Et svar til Bjoerg Svanes. *Maal og Minne 1–2*, 1–18.

Stemmer, Brigitte (1981). Kohaesion im gesprochenen Diskurs deutscher Lerner des Englischen. *Manuskripte zur Sprachlehrforschung 16*. Bochum, Ruhr-Universitaet.

Svanes, Bjoerg (in press). En undersoekelse av realisasjonsmoenstret for spraakhandlingen "aa be noen om aa gjoere noen". *Maal og Minne 3–4*.

Takahashi, Satomi (1987). *A contrastive study of indirectness exemplified in L1 directive speech acts performed by Americans and Japanese*. Unpublished master's thesis, University of Illinois at Urbana-Champaign.

Takahashi, Satomi (forthcoming). *Exploring comprehension processes of nonliteral utterances and some implications for automaticity*. Ms., University of Hawaii at Manoa.

Takahashi, Satomi & DuFon, Peggy (1989). *Cross-linguistic influence in indirectness: The case of English directives performed by native Japanese speakers*. Unpublished manuscript, University of Hawaii at Manoa.

Takahashi, Tomoko & Beebe, Leslie M. (1987). The development of pragmatic competence by Japanese learners of English. *JALT Journal 8*, 131–155.

Tanaka, Noriko (1988). Politeness: Some problems for Japanese speakers of

English. *JALT Journal 9*, 81–102.

Tanaka, Shigenori & Kawade, Saiki (1982). Politeness strategies and second language acquisition. *Studies in Second Language Acquisition 5*, 18–33.

Trosborg, Anna (1987). Apology strategies in natives/non-natives. *Journal of Pragmatics 11*, 147–167.

Walters, Joel (1979). The perception of deference in English and Spanish. In Carlos A. Yorio, Kyle Perkins & Jacquelyn Schachter (Eds.) *On TESOL '79* (pp. 288–296). Washington, D.C.: TESOL.

Walters, Joel (1980). Grammar, meaning and sociocultural appropriateness in second language acquisition. *Canadian Journal of Psychology 34*, 337–345.

Wolf, R. M. (1988). Questionnaires. In John P. Keeves (Ed.) *Educational Research, methodology, and measurement* (pp. 478–482). Oxford: Pergamon.

Wolfson, Nessa (1989a). *Perspectives. Sociolinguistics and TESOL*. New York: Newbury House.

Wolfson, Nessa (1989b). The social dynamics of native and nonnative variation in complimenting behavior. In Miriam Eisenstein (Ed.), *The dynamic interlanguage* (pp. 219–236). New York:Plenum.